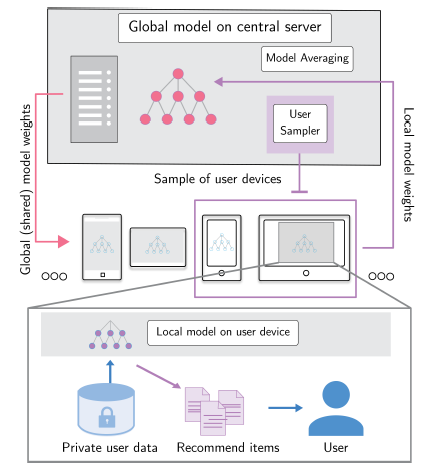


ACM SIGKDD 2020

联邦学习过程：



FedAvg应用到推荐系统的问题：

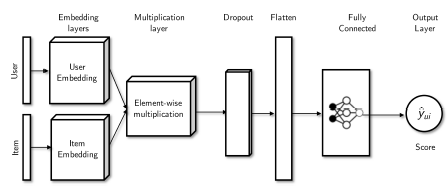
1. 推荐效果差
2. 收敛速度慢
3. 由上两点导致的通信轮次过多和客户端通信资源浪费

本文目标：设计一个用于推荐系统的联邦算法，以获得更准确地推荐、不浪费不必要的开销、降低通信成本。

限制：

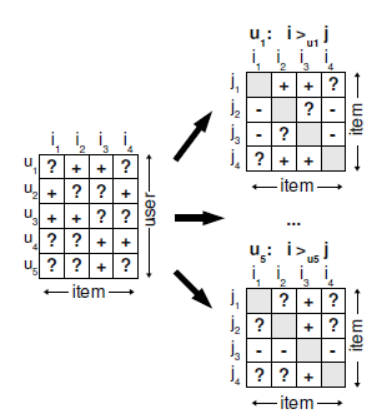
推荐系统的参数应当尽量少（反例：NCF），并且是经过嵌入后得到的推荐系统。

例1：GMF网络（Loss：pointwise-MSE pairwise-交叉熵）：



例2：BPR（贝叶斯个性化排序）

假设不同用户间的偏好没有关系，同一用户对两个物品的偏好和其它物品无关。



求最大似然参数：，。梯度上升或拟牛顿法。

选择用户：FedAvg是随机选择用户

ActvSAMP：

通过client的metadata（标签），确保匿名性

标签：a set of summary statistics (e.g. mean and entropy) of their user profiles.

方法：k-means

轮询k个集合，直到m个设备。

可能的改进：Other privacy-preserving features (e.g. region, device type)

聚合模型：FedAvg对被选用户求平均

ActvAGG：

1.全局更新参数（这些参数与嵌入无关）

2.item嵌入：对于每一个item嵌入的w[i]，根据前一轮次和这一轮次的差作为更新权重，用于求嵌入的delegate（指客户端client）的平均值，将该值作为这些delegate的新嵌入

3.重新划分所有的样本

3.对每个sample，找到所在类，对属于这些类的全部样本，但不在本次采样的样本：

将该样本的前一轮次和这一轮次的差作为更新权重。

4.对于图中的每个类，且非采样样本，更新的方式是在上一轮次更新的基础上，加上刚才的累积更新权重的均值，再乘一个系数

复杂度：

, where n is the number of users to be clustered, k is the number of clusters, and d is

the size of the dimensions or embedding size.

优化：Lloyd’s algorithm where is the number of iterations needed until convergence)

实验结果：

GMF model：

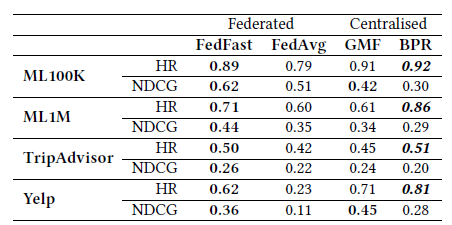
数据集：Movie Lens 100K

21,009 parameters 84.5 kB weights embedding size: 8

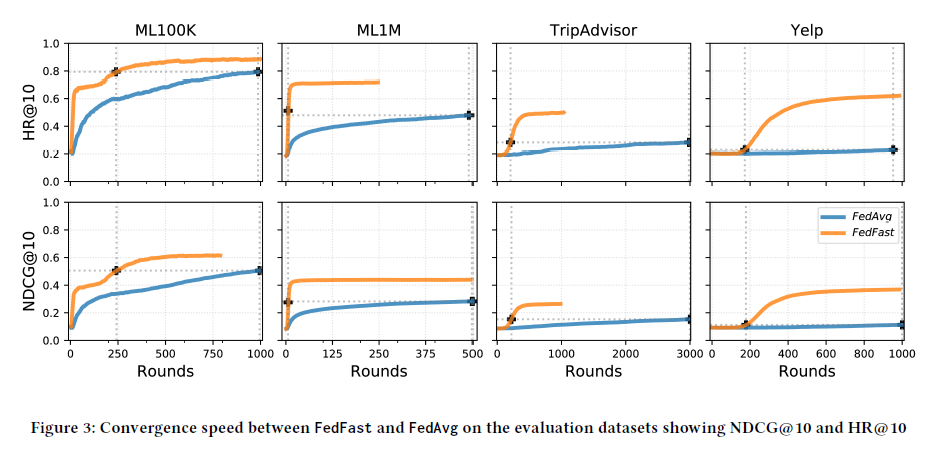
23,635 parameters(+12.4%) embedding size: 9 Hit Ratio (+0.6%)

考虑到移动设备的处理性能和存储空间，及上述分析，采用小规模的嵌入更加合理。

性能比较（前10个）：t-test p<0.01

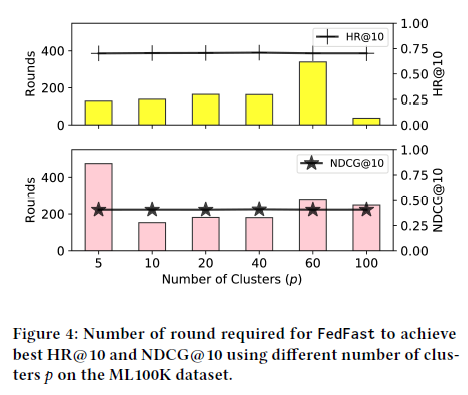


收敛性：

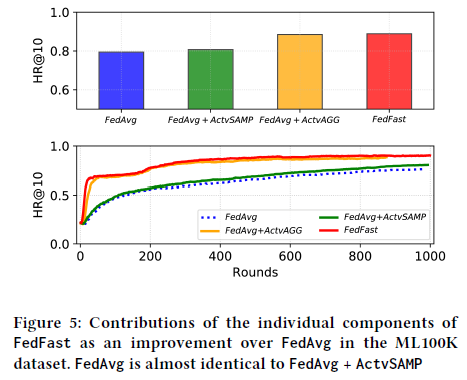


在某一点后，每一轮都比FedAvg好，由此可见这是一个anytime的算法。

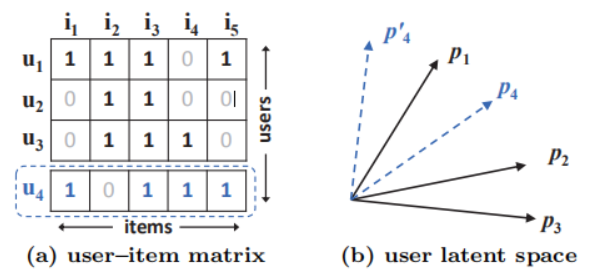
收敛轮次：

轮次少

个体比较：

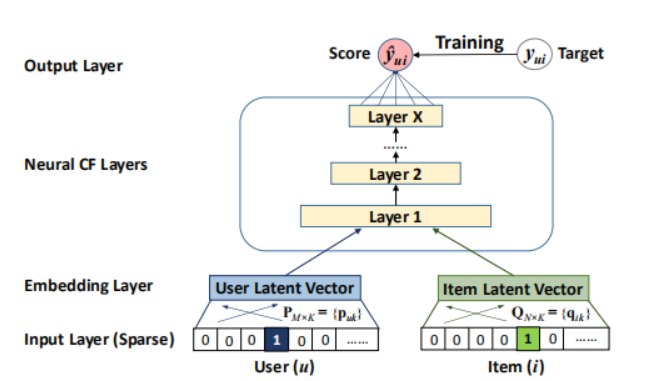


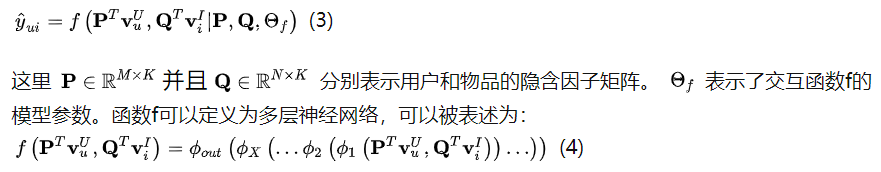
MF的局限性：



（a）中使用Jaccard作为相似性度量（交集/并集），（b）中使用cos作为相似性度量。

NCF结构：





Q：为什么MF（即前文中提到的GMF）是特殊的NCF（给出映射）？

推荐系统指标：

Hit Ratio (HR)：命中率

Normalized Discounted Cumulative Gain (NDCG)：

给定相关性（即结果是否相关）：[1,2,3]

累计增益（CG）：1+2+3

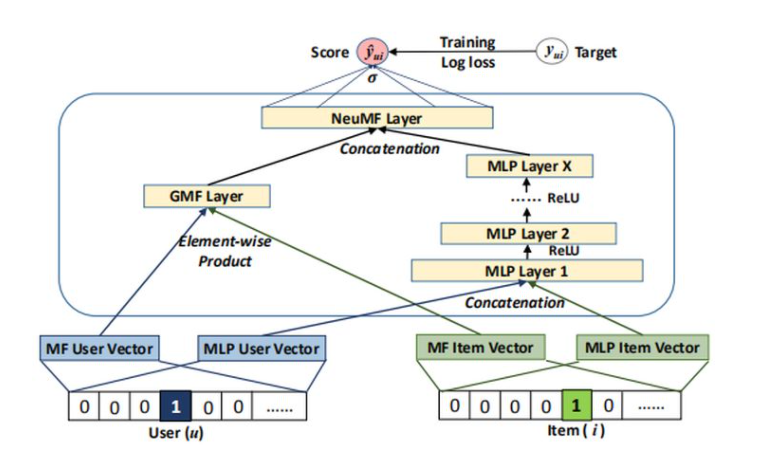
折损累计增益（DCG）：排名i越靠后，越不重要



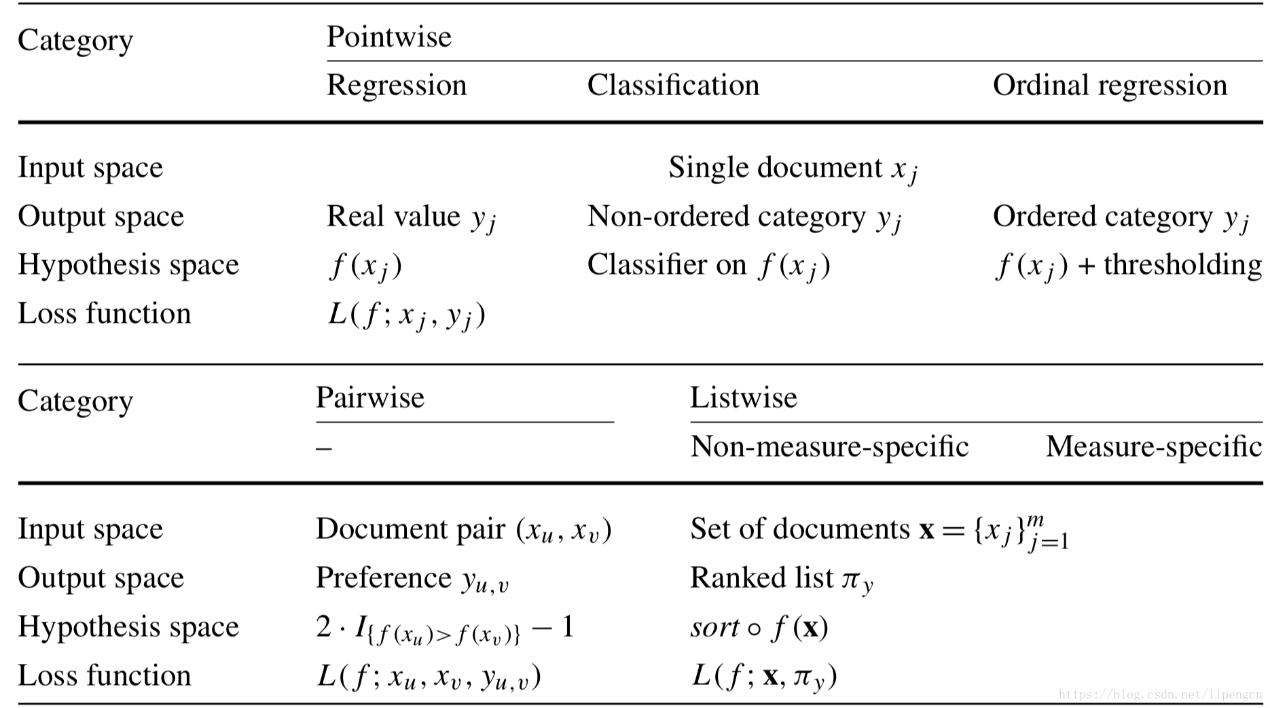
NDCG=DCG/IDCG

其中，IDCG指在最正确情况下（上例则为[3,2,1]）的DCG值。

NeuMF模型（NCF+MLP）：



RS中的loss functions：



领域：Github+ML App Store+Developers RS+FL